Research report 154

MULTIRESOLUTION IMAGE SEGMENTATION COMBINING REGION AND BOUNDARY INFORMATION

A Bhalerao
R Wilson

(RR154)

A new approach to image segmentation is presented that integrates region and boundary information within a great framework of Maximum a Posteriori (MAP) estimation and decision theory. The algorithm employs iterative, decision-directed estimation performed on a spatially localised basis but within a multiresolution representation. The use of a multiresolution technique ensures both robustness in noise and efficiency of computation, while the model-based estimation and decision process is both flexible and spatially local, thus avoiding assumptions about global homogeneity or size and number of regions. The method gives accurate segmentations at low signal-to-noise ratios and is shown to be more effective than previous methods in capturing complex region shapes.

Department of Computer Science
University of Warwick
Coventry CV4 7AL
United Kingdom

April 1990

Contents

List of Figures

List of Tables

Multiresolution Image Segmentation Combining Region and Boundary Information

A. Bhalerao R. G. Wilson

Abstract

A new approach to image segmentation is presented that integrates region and boundary information within a general framework of Maximum *a Posteriori* (MAP) estimation and decision theory. The algorithm employs iterative, decision-directed estimation performed on a spatially localised basis but within a multiresolution representation. The use of a multiresolution technique ensures both robustness in noise and efficiency of computation, while the model-based estimation and decision process is both flexible and spatially local, thus avoiding assumptions about global homogeneity or size and number of regions. The method gives accurate segmentations at low signal-to-noise ratios and is shown to be more effective than previous methods in capturing complex region shapes.

[0]This work is supported by SERC and Shell Research Ltd.

1 Introduction

Over recent years there has been a growing interest in the use of multiresolution, or scale space, methods in the segmentation of images. Such methods can broadly be classed as either region-based or edge-based. Among the former, notable examples were the 'split-and-merge' algorithm of Chen and Pavlidis [1], and the linked pyramid methods of Burt *et al* [2]. More recently, Spann and Wilson described a quad-tree based algorithm employing classification at a low spatial resolution combined with downward directed boundary estimation, for the segmentation of images containing arbitrary numbers of regions of homogeneous grey level or texture [3], [4]. A further development based on adaptive thresholding was reported by Spann and Horne [5]. An attempt to place such methods in the more precise statistical framework of Markov Random Fields (MRF) was reported recently by Gidas [6].

Because of the inherent limitations of region-based methods, a number of authors have preferred edge-based techniques. Among these, notable early work was reported by Thurston and Rosenfeld [7] and Marr [8]. More recently, the 'optimal' edge detector of Canny [9] and the 'edge-focusing' technique of Bergholm [10] have appeared in the literature, as well as methods based on spatial frequency analysis, such as that of Calway and Wilson [11].

While either methodology can prove satisfactory under certain conditions, it would clearly be beneficial to combine region and edge information in a single structure to make maximum use of all the available information relevant to a given segmentation. Such approaches have recently been described in [12], [13]. It is the purpose of this paper to describe a new method of integrating region and boundary information within a general framework of Maximum *a Posteriori* (MAP) estimation and decision theory (e.g. [12], [14]). The algorithm employs iterative, decision-directed estimation performed on a spatially localised basis but within a multiresolution framework. The use of a multiresolution technique ensures both robustness in noise and efficiency of computation, while the model-based estimation and decision process is both flexible and spatially local, thus avoiding assumptions about global homogeneity or size and number of regions which characterise some of the earlier algorithms. A description of the algorithm is followed by some results on the segmentation of objects with complex shapes against a high level of noise ($<$ 0dB SNR) and on a natural scene. The results show that the new technique promises to be more effective than previous methods in capturing complex region shapes.

2 The Algorithm

2.1 General Structure

The overall structure of the algorithm is depicted in Figure 1. It has three main components:

(i) Multiresolution tessellation

The first processing step is to produce an inhomogeneous tessellation into a set of square regions, each of which corresponds to a unique node in a quad-tree (Fig. 4(c)). The decision process producing this tessellation is a bottom-up procedure based on an estimate of the likelihood that a given region is homogeneous under some appropriate model of region properties (e.g. for grey level, colour or texture). The aim is to produce a representation with as few regions as possible, subject to the homogeneity constraint.

(ii) Region growing

Nodes in the resulting tessellation are then grouped using a combination of MAP decision process and decision directed estimation, within a region adjacency graph. This process is aimed at producing a set of connected regions, each having well defined boundary and region average property (Fig. 5(b) and (c)).

(iii) Boundary estimation

The coarse boundary estimate from the region growing process is then refined using a separately computed estimate of local boundary orientation and position. This process is a form of 'hill-climbing' technique which seeks the best positions to place a set of vertices defining the boundary, based on both boundary energy and region estimation errors.

The overall computational paradigm is essentially that of producing a 'least-cost' or minimum m.s.e fit to the data, within the constraint of a model of region and boundary properties. In this respect, it shares some features with the Bayesian methods described in [12] and [13]. However, at this stage it has not proved necessary to resort to stochastic relaxation, as deterministic procedures have produced satisfactory results.

2.2 Quad-Tree Smoothing and Node Selection

The first step is a smoothing operation performed using a quad-tree [15].

Let $d(i,j), 0 \leq i,j, < N$, be the original image size $(N \times N), N = 2^m$. The quad tree $q(i,j,l), 0 \leq l \leq m$, where l represents the level, is formed by:

$$\begin{aligned} q(i,j,l) &= 0.25[q(2i,2j,l-1) + q(2i+1,2j,l-1) \\ &\quad + q(2i,2j+1,l-1) \\ &\quad + q(2i+1,2j+1,l-1)], \; l > 0 \\ q(i,j,0) &= d(i,j) \end{aligned} \tag{1}$$

The quad-tree is a fast smoothing operation that trades off noise reduction by smoothing against spatial resolution (Fig. 4(a)). Quad-tree nodes are selected based on parent child variances $v(i,j,l), l > 1$ where:

$$\begin{aligned} v(i,j,l) &= 0.25[(q(2i,2j,l-1) - q(i,j,l))^2 \\ &+ ((q(2i+1,2j,l-1) - q(i,j,l))^2 \\ &+ ((q(2i,2j+1,l-1) - q(i,j,l))^2 \\ &+ ((q(2i+1,2j+1,l-1) - q(i,j,l))^2] \end{aligned} \tag{2}$$

and the sample variance for a given level is just the average

$$\overline{v}_l = \frac{1}{(2^{m-l})^2} \sum_{0 \leq i,j < 2^{m-l}} v(i,j,l) \tag{3}$$

If $v(i,j,l) > \alpha \overline{v}_{l-1}$ then the associated node $q(i,j,l)$ is marked. This process is performed bottom-up, and when complete the lowest marked nodes are selected to represent a square block $(2^l \times 2^l)$ of the image. The assumption is that $\overline{v}_{l-1}$ is a good estimator of the image 'within region' variance on level l of the quad-tree so that any variation in a 2×2 block which is significantly greater than this will be due to the node $q(i,j,l)$ representing more than one region. The factor α is used to control the extent of the variability allowed.

This selection procedure effects a tessellation of the original image into square blocks of different sizes, where a selected node $q(i,j,l)$ represents the image block $\{d(x,y), (2i,2j) \leq (x,y) < (2i+2^l, 2j+2^l)\}$ (Fig. 4(b)(c)).

2.3 Region Adjacency Graph

A graph $R = \{r_1, r_2, \ldots, r_{n_r}\}$ which consists the set of n_r selected nodes $q(i,j,l)$, with each $r_k \in R$ being associated with a selected quad-tree node: $r_k = q(i,j,l)$ for some i,j,l. A first order neighbourhood ∂_k for a node $r_k \in R$ can be defined as being all nodes that represent image blocks bordering the image block represented by node r_k (Fig. 2). The simplest case is when all the neighbours are from the same level as r_k, in which case the neighbourhood is the 4-neighbours of r_k. This structure is a modification of the conventional region adjacency graphs used in segmentation [16] which takes advantage of the scale invariance and computational properties of the quad-tree tessellation.

2.4 Iterated Decision-directed Estimation

This part of the processing is based on a normal model for the data, in which regions are assumed to have independently selected means. Each node $r_i \in R$ is iteratively averaged with its neighbours in the neighbourhood ∂_i:

$$r_i^n = \frac{1}{N(\partial_i)} \sum_{j \in \partial_i} s_{ij}^n r_j^{n-1} \tag{4}$$

where $N(\partial_i)$ is the number of neighbours of r_i and n is the iteration number. s_{ij}^n is a switching function determined by the MAP test that a given link is 'on', given the data at the two nodes:

$$P_{kl}(L)P(r_i, r_j \mid L) > P_{kl}(\overline{L})P(r_i, r_j \mid \overline{L}) \tag{5}$$

where $P(r_i, r_j \mid L)$ is the conditional probability density of the data assuming that nodes i and j, associated with quad tree nodes from levels k and l respectively, are connected (L = True) and therefore have the same mean.

$$\begin{aligned} P(r_i, r_j \mid L) &\propto \exp\left(\frac{-(r_i^n - r_j^n)^2}{2(\overline{v}_k + \overline{v}_l)}\right) \\ P(r_i, r_j \mid \overline{L}) &\propto \exp\left(\frac{-(r_i^n - r_i^0)^2}{2\overline{v}_k}\right) \\ &\quad \cdot \exp\left(\frac{-(r_j^n - r_j^0)^2}{2\overline{v}_l}\right) \end{aligned} \tag{6}$$

After a decision is made about each link from the node r_i, its colouring is then taken as the minimum mean square error (MMSE) estimate of the associated 'on' neighbours.

An appropriate estimate of $P_{kl}(L)$ can be based on the ratio:

$$\frac{\text{number of edge blocks}}{\text{total number of blocks}} \tag{7}$$

This, of course, is dependent on knowing the number and size of the regions in the image. In practice, it is possible to use heuristic arguments, perhaps supplemented by *a priori* knowledge of region sizes, to form an initial estimate of $P_{kl}(L)$. A better estimate is then derived from the data, before the first iteration, using the switching function state (equation 5).

The convergence criterion for the averaging is when there is no change in the state of any link i.e. that $s_{ij}^n = s_{ij}^{n-1}$ for all i, j. This is typically achieved in 10-20 iterations, (Fig. 5(a)).

2.5 Boundary Graph

After convergence a set of boundary nodes $B = \{b_1, b_2, \ldots, b_{n_b}\}$ are sited, one for each link that is off on the convergent iteration. The node b_i is positioned at the mid-point of the block boundary shared by the region nodes linked by the off link. Adjacent boundary nodes are then linked together; boundary nodes are adjacent if they lie on adjacent sides of the same block (Fig. 3). Linking together these boundary nodes creates a *dual* of the region adjacency graph (Fig. 5(b)).

A corridor is defined along the edges of the boundary graph which is used to mask out data for the orientation estimation and enhancement process. The width of this corridor is a multiple of the smallest block size in the region tessellation.

2.6 Orientation Estimation and Enhancement

A local orientation estimation is performed within the corridor defined by the boundary obtained from the block segmentation. The orientation estimate is performed on a Gaussian pyramid of the original image [17]. Such a pyramid overcomes the inherent block aliasing of the 2×2 quad-tree averaging process and therefore yields better orientation estimates. The orientation estimator used is a fast, 4-point operation [18], that produces a vector at each point. Given the 4-pixels in a block a, b, c, d the estimate of the local orientation θ can be expressed as a vector:

$$\underline{v}_{2\theta} = \begin{pmatrix} 2(b-c)(d-a) \\ (b-c)^2 - (d-a)^2 \end{pmatrix} \tag{8}$$

The magnitude of this vector is a measure of the local edge energy and the argument *twice* the local orientation. The doubling of the the orientation, originally proposed by Knutsson [19], is essential to make operations such as averaging of the vector orientations meaningful.

Since the estimator is a gradient operator, it is very sensitive to noise, (Fig. 5(c)). At higher levels of the pyramid, however, where the noise has been smoothed out, the estimate is more reliable, again at the loss of spatial resolution. The enhancement procedure utilises the more reliable information at higher levels by propagating it down the pyramid to restore the estimate at lower levels, thus regaining spatial resolution. The procedure is a variation of the minimum mean square error (MMSE) estimator developed by Clippingdale [20].

The enhancement is a recursive, two stage process: a 'vertical' propagation of information from parents to children is followed by an anisotropic smoothing of the propagated data. Both stages use linear MMSE estimation to obtain a constrained optimal estimate, where the constraints ensure that the computational burden is not excessive.

2.6.1 Vertical Propagation

An estimation $\hat{\mathbf{s}}_c$ of the orientation vector of a node on the child level c can be defined as a linear combination of the data $\mathbf{x}_c$ and the estimate of the parent $\mathbf{s}_p$:

$$\hat{\mathbf{s}}_c = \alpha \mathbf{x}_c + \beta \hat{\mathbf{s}}_p \tag{9}$$

$$\mathbf{x}_c = \mathbf{s}_c + \mathbf{n}_c \tag{10}$$

where $\mathbf{n}_c$ is the additive noise on the child level. By the orthogonality principle:

$$E(\mathbf{s}_c - \hat{\mathbf{s}}_c)\mathbf{x}_c = 0 \tag{11}$$

$$E(\mathbf{s}_c - \hat{\mathbf{s}}_c)\hat{\mathbf{s}}_p = 0 \tag{12}$$

The noise variance on level c can be estimated from the measured parent-child variances of level 0. Assuming that the signal and noise are uncorrelated across and between levels of the orientation pyramid it is possible to solve the simultaneous equations 11 and 12 for the feedback coefficients α and β.

This vertical propagation is both cheap computationally and leads to a significant reduction in the m.s.e of the vector components. Note that the additive feedback of the orientation vectors is only possible if they are in the double angle representation.

2.6.2 Anisotropic Filtering

The propagated orientation is filtered anisotropically by a 2-d Gaussian filter $h(x, y, \theta)$:

$$h(x, y, \theta) = h_x(x_\theta) h_y(y_\theta) \tag{13}$$

$$\begin{aligned} h_x(x) &= \alpha_x e^{-\beta_x x^2} \\ h_y(y) &= \alpha_y e^{-\beta_y y^2} \end{aligned} \tag{14}$$

where θ is the local orientation at the point (x, y) and:

$$x_\theta = x\cos(\theta) - y\sin(\theta) \tag{15}$$

$$y_\theta = x\sin(\theta) + y\cos(\theta) \tag{16}$$

The parameters that determine the filter shape $(\alpha_x, \beta_x, \alpha_y, \beta_y)$ are derived by posing the problem as a linear minimum mean squares estimation. The correlation function of the orientation data is assumed to be separable in the orientation of the feature and like the filter to be used, Gaussian in shape, because it is primarily the result of the smoothing operation of the Gaussian pyramid.

$$R_{ss}(x, y, \theta) = R_{ss_x}(x_\theta) R_{ss_y}(y_\theta) \tag{17}$$

$$\begin{aligned} R_{ss_x} &= e^{-\rho_x x^2} \\ R_{ss_y} &= e^{-\rho_y y^2} \end{aligned} \tag{18}$$

The noise correlation is similarly separated into:

$$\begin{aligned} R_{nn_x}(x) &= \sigma\delta(x) \\ R_{nn_y}(y) &= \sigma\delta(y) \end{aligned} \tag{19}$$

The orientation energy will be highly correlated along the direction of the feature, with spread in the perpendicular direction being greater at lower levels of the orientation pyramid. The 1-d problem can now be stated as:

$$e = E\left[\mathrm{s}(x) - \int_{-\infty}^{\infty} \alpha_x e^{-\beta_x x^2} e^{-\rho_x x^2} dx\right]^2 \tag{20}$$

Minimising e w.r.t. β_x:

$$\begin{aligned} \frac{\partial e}{\partial \beta_x} &= -2\int_{-\infty}^{\infty} \alpha_x e^{-\beta_x x^2} e^{-\rho_x x^2} dx \\ &+ \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \alpha_x^2 e^{-\beta_x x^2} e^{-\beta_x w^2} \\ &\left[e^{-\rho_x (x-w)^2} + \sigma\delta(x-w)\right] dx\, dw \end{aligned} \tag{21}$$

After some manipulation it can be shown that:

$$\beta_x = \frac{\rho_x + 2\rho_x^2\sqrt{2\pi}}{\sigma + (\beta_x + 2\rho_x)^{3/2}} \tag{22}$$

$$\alpha_x = \sqrt{\frac{2\beta_x(\beta_x + 2\rho_x)}{(\beta_x + \rho_x)}} \frac{1}{(\sqrt{2\pi} + \sigma\sqrt{\beta_x + 2\rho_x})} \tag{23}$$

By least squares fitting a Gaussian function to the correlation data $R_{ss}(x, y)$ it is possible to determine ρ_x and ρ_y. Equation 22 can then be solved numerically and hence α_x and α_y giving the MMSE filter $h(x, y, \theta)$.

The filtering is performed by iterated 5×5 spatial convolution of a set of N_f oriented filters $h(x, y, \theta_k)$ where $\theta_k = \frac{k\pi}{N_f - 1}$ $0 \leq k \leq N_f$. The appropriate filter is selected at each point by the argument θ of the downward propagated orientation estimate. This procedure is not in practice computationally expensive as it is only performed on the masked region obtained from the initial boundary estimate.

Table 1 shows results of the enhancement process on the *shapes* image (Fig. 4(a)) from the starting level (4) to full spatial resolution, level 0. The first column shows a correlation measure that compares the orientation estimates of the input and output against an orientation estimate from a noise free version of the original image. This measure is defined as:

$$c_l = \frac{\sum_{i,j\in\mathcal{R}_l} |\mathrm{s}(i,j,k)||\hat{\mathrm{s}}(i,j,l)| \cos[\angle(\hat{\mathrm{s}}(i,j,k)) - \angle(\mathrm{s}(i,j,l))]}{\sum_{i,j\in\mathcal{R}_l} |\mathrm{s}(i,j,k)||\hat{\mathrm{s}}(i,j,l)|} \tag{24}$$

where $\mathcal{R}_l$ is the set of points of the corridor defined along the edges of the initial boundary graph on level l (Section 2.5). The numerator of c_l is the inner product of the estimate and reference orientation vectors, which will be low for vectors that deviate significantly in angle from the reference. The second column on Table 1 shows the angle estimation error in degrees which is calculated from the correlation measure by:

$$D_l = 0.5 \cdot \cos^{-1}(c_l) \tag{25}$$

Note that all angles are double angles, hence the 0.5 factor in equation 25. It can be seen from Table 1 that the enhancement process more than halves the average error at the highest resolution.

Figures 5(d) shows the magnitude of the enhanced orientation estimate of the *shapes* image shown in Figure 5(c).

2.7 Boundary Relaxation

The enhanced orientation estimate is now used to control the relaxation of the boundary graph B. Each boundary node b_i is moved to maximise a local cost function based on the orientation and grey level statistics gathered in the vicinity of the node. The method used has similarities to the boundary refinement algorithms used in [1], [3] and also related to the contour modification process outlined by Pavlidis and Liow [21], [22].

Each boundary node b_i is permitted to move either perpendicular or parallel to the average orientation along the links from b_i. The cost function is defined as:

$$C(b_i) = \frac{1}{N(\partial_i)} \left[A \sum_{j \in \partial_i} e(\phi(l_{ij})) + B \sum_{j \in \partial_i} p(l_{ij}) \right] \tag{26}$$

where ∂_i is the neighbourhood of the node b_i and l_{ij} is the link between b_i and its neighbour b_j. The energy function $e(\phi(l_{ij}))$ is the projection of the orientation energy of the vectors along the link onto the orientation $\phi(\cdot)$ of the link l_{ij}. $p(l_{ij})$ is derived from the image statistics in rectangular regions either side of the link l_{ij}, which represents the likelihood that the region is split by the link:

$$p(l_{ij}) = \frac{(m_1 - m_2)^2}{(\sigma_1^2/n_1) + (\sigma_2^2/n_2)} \tag{27}$$

where (m_1, m_2), (σ_1^2, σ_2^2), (n_1, n_2) are the means, variances and populations of the rectangular regions either side of the link l_{ij}. The coefficients A and B control the contribution of the two energies toward the merit function. During perpendicular movement $A = B = 1.0$, where as during parallel movement $A = 1.0$, $B = 0.0$.

All boundary nodes are initially forced to move perpendicular to the local orientation; when a node ceases to move it is allowed to move in the parallel direction, away from the neighbourhood link which has the largest orientation energy $e(\cdot)$.

One iteration of the relaxation process consists of finding the maximum of the cost function $C(\cdot)$ for all nodes by searching in the boundary corridor defined above. After each iteration there is a merging step to merge any nodes that are within some minimum distance set to be the size of the smallest block of the region tessellation. The boundary relaxation, and hence the segmentation, is complete when there is no significant average movement of the boundary nodes.

Figures 6(a) and (b) show the relaxation process after 1 and 4 iterations respectively. Each boundary node on these images is marked with a vector showing the direction of its movement during the iteration with Figure 6(a) showing perpendicular movement of boundary nodes and Figure 6(b) showing their movement along the boundary. Figure 6(c) shows the completed segmentation after 7 iterations.

3 Experimental Results

Figures 6(c) and (d) show segmentation results on a synthetic image of size 256×256, (8-bit pixel), to which Gaussian noise has been added with a standard deviation $\sigma = 25$, (Fig 4(a)). The luminance difference between the object and the background is 20 giving an inter-region signal to noise ratio [5] of $20/\sigma = 0.8$. The segmentation of the circle compares well with the results obtained by Spann and Wilson [3]. Its performance on the objects that have sharp corner features, however, is superior to those reported elsewhere.

The results on the natural *boats* image, also 256×256, 8-bit, are shown in Figures 8(c) and (d). The segmentation is not perfect because the underlying image model does not take sufficient account of luminance variation across regions. Also the boundary relaxation does not place restrictions to stop the collapsing of long thin regions, such as the masts of the boat, into regions of zero width.

4 Conclusions

A new algorithm for image segmentation, having a number of interesting features, has been described and been shown to be effective in locating regions and their boundaries with high accuracy from noisy data. In applications where region shape is important, it may have a number of advantages over previously reported methods. It is computationally fast, flexible with respect to region and boundary models and uses only local processing.

There are several areas, however, in which consolidation and extension are required. The extension from grey level to other region properties is a simple one (e.g. [4]) as is that from 2-D to problems in higher dimensions. However, there is a need to replace the simple estimation and decision models (section 2.4) with more appropriate models, to deal with more realistic problems. This is not a problem in principle, since

the statistical models on which decisions and estimates are based is a local one. Similarly, if 'line' type structures, such as the masts of the *boats* in Figure 7(a), are to be segmented, the region model will have to be significantly modified. Lastly, a more precise statistical model of the region boundaries should lead to better estimates. Such developments are currently under way.

Acknowledgement

This work is supported by the U.K. S.E.R.C. and by Shell Research Ltd.

References

[1] P. C. Chen and T. Pavlidis. Image segmentation as an estimation problem. In A. Rosenfeld, editor, *Image Modelling*, pages 9–28. Academic Press, 1981.

[2] T. H. Hong P. J. Burt and A. Rosenfeld. Segmentation and estimation of image region properties through cooperative hierarchical computation. *IEEE Trans. Sys, Man Cyber.*, 11(12):802–809, 1981.

[3] M. Spann and R. G. Wilson. A quad-tree approach to image segmentation which combines statistical and spatial information. *Pattern Recognition*, 18(3/4):257–269, 1985.

[4] Roland G. Wilson and M. Spann. *The Uncertainty Principle in Image Processing.* Pattern Recognition and Image Processing Series. Research Studies Press Ltd, 1988.

[5] M. Spann and C. Horne. Image segmentation using a dynamic thresholding pyramid. *Pattern Recognition*, 22(6):729–732, 1989.

[6] B. Gidas. A renormalization group approach to image processing problems. *IEEE Trans. P.A.M.I.*, 11(2):164–180, 1989.

[7] A. Rosenfeld and M. Thurston. Edge and curve detection for visual scene analysis. *IEEE Trans. Comp.*, 5(20):562–569, 1971.

[8] D. Marr and E. Hildreth. Theory of edge detection. *Proceedings of the Royal Society London*, B207:187–217, 1980.

[9] J. Canny. A computation approach to edge detection. *IEEE Trans. P.A.M.I.*, 8:679–698, 1986.

[10] F. Bergholm. Edge focusing. *IEEE Trans. P.A.M.I.*, 9(6):726–741, 1987.

[11] A. Calway and R. G. Wilson. A unified approach to feature extraction based on an invertible image transform. In *Proc. 3rd IEE Int. Conf. Image Processing*, pages 651–655, Warwick, U.K, 1989.

[12] J. Besag. On the statistical analysis of dirty pictures. *Journal of the Royal Statistical Society*, 48(3):259–302, 1986.

[13] S. Geman and D. Geman. Stochastic relaxtion, Gibbs distributions, and the Bayesian restoration of images. *IEEE Trans. P.A.M.I.*, 6(6):721–741, 1984.

[14] C. W. Therrien. *Decision Estimation and Classification.* J. Wiley and Sons, 1989.

[15] S. L. Tanimoto and T. Pavlidis. A hierarchical data structure for picture processing. *Comp. Vision Graphics and Im. Proc.*, (4):104–119, 1975.

[16] S. Levialdi. Basic ideas for image segmentation. In O. D. Faugeras, editor, *Fundamentals in computer vision*, pages 239–261. Cambridge University Press, 1983.

[17] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. Comp.*, COM-31:532–540, 1983.

[18] M. Todd. *Image Data Compression Based on a Multiresolution Signal Model.* PhD thesis, Department of Computer Science, The University of Warwick, UK, November 1989.

[19] H. Knutsson. *Filtering and Reconstruction in Image Processing.* PhD thesis, University of Linköping, Sweden, 1982.

[20] S. Clippingdale. *Multiresolution Image Modelling and Estimation.* PhD thesis, Department of Computer Science, The University of Warwick, UK, September 1988.

[21] T. Pavlidis and Y. T. Liow. Integrating region growing and edge detection. *IEEE Trans. P.A.M.I.*, 12(3):225–233, 1990.

[22] Y. T. Liow and T. Pavlidis. Use of shadows for extracting buildings in aerial images. *Comp. Vision Graphics and Im. Proc.*, (49):242–277, 1990.

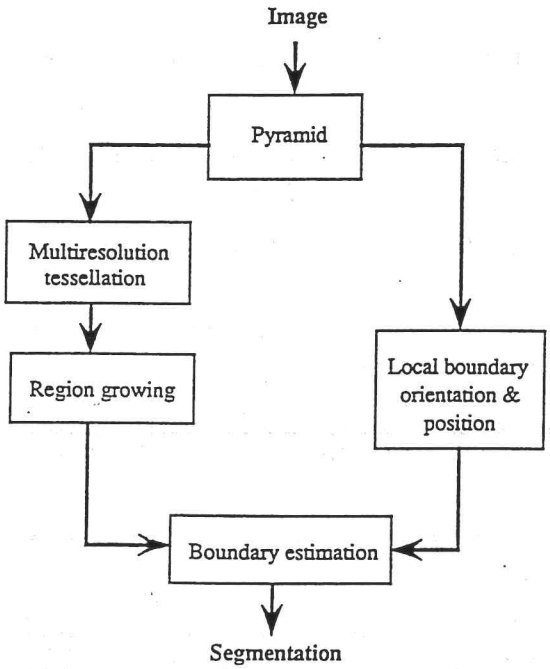

Figure 1: Algorithm structure.

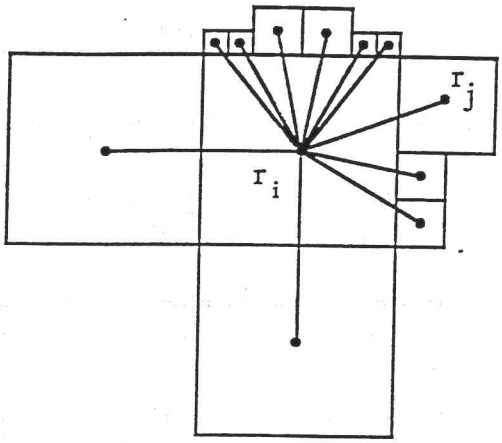

Figure 2: Region adjacency.

level	*Input*		*Output*	
	c_l	D_l (°)	c_l	D_l (°)
0	0.75	20.7	0.94	10.3
1	0.84	16.6	0.94	9.9
2	0.94	10.4	0.96	8.0
3	0.98	6.3	0.98	5.6
4	0.99	3.4	0.99	3.4

Table 1: Orientation enhancement results on *shapes* image.

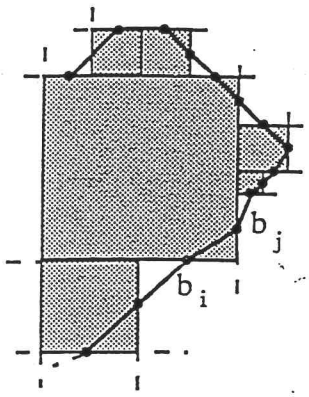

Figure 3: Placement of boundary nodes.

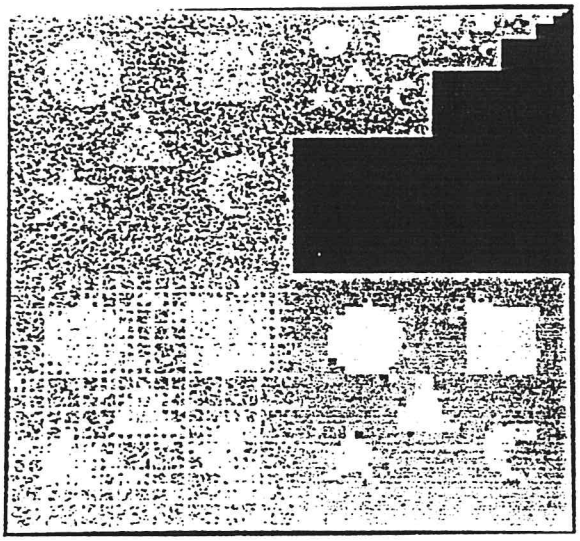

Figure 4: (a) Quad-tree of original 256 × 256, *shapes* image, $\sigma = 25$. (b) Block tessellation. (c) Quad-tree means of selected nodes.

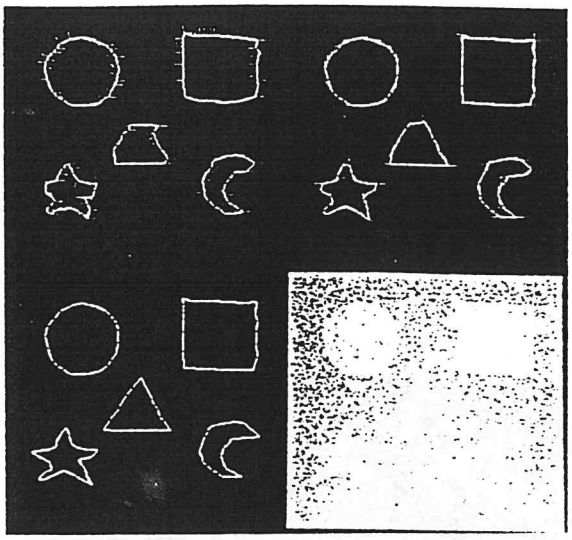

Figure 6: (a) Boundary relaxation after 1, (b) 4 and (c) 7 iterations. (d) Segmentation result superimposed on original.

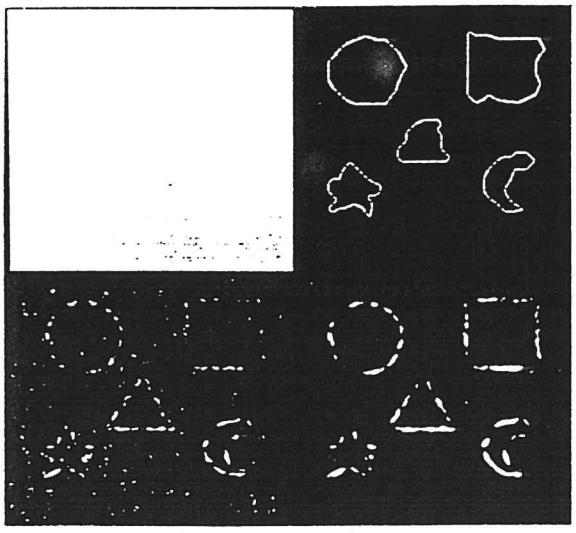

Figure 5: (a) Converged region estimation (10 iterations). (b) Initial boundary estimate. (c) Orientation estimate (magnitude). (d) Enhanced orientation estimate (magnitude).

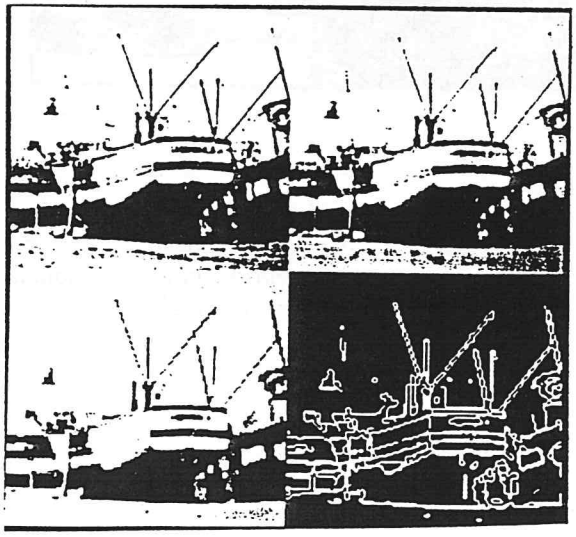

Figure 7: (a) Original 256 × 256, *boats* image. (b) Quad-tree means of selected nodes. (c) Converged region estimation (8 iterations). (d) Initial boundary estimate.

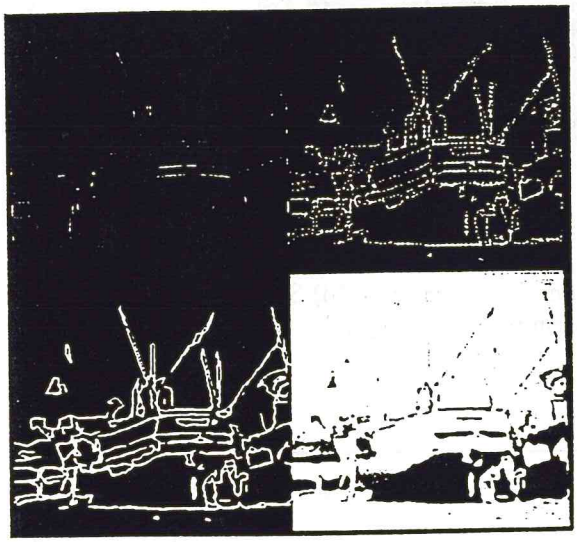

Figure 8: (a) Orientation estimate (magnitude), *boats* image. (b) Boundary relaxation after 1 iteration. (c) Converged relaxation (5 iterations). (d) Superimposed result on original.